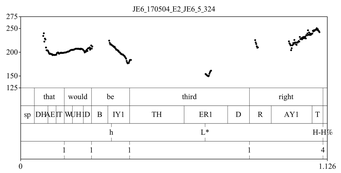
Spurious Pitch Movements in American English Polar Questions

PRINCETON UNIVERSITY UCLA

Byron Ahn bta@princeton.edu **Z.L. Zhou** zlzhou@ucla.edu

Initial Observation

- American English (AmE) Polar Questions (PQs) regularly exhibit pitch movements upward before a low accent



- These are **Spurious Pitch Movements (SPMs)** under an MAE_ToBI model
  - Pitch accent **alignment**:
    - Pitch accents are aligned with intonationally (=post-lexically) prominent syllables
    - Leading/Trailing tones are aligned to be within one syllable of the prominent syllable
  - Cover tone **directionality**:
    - AmE cover tones are aligned with the end of an intermediate phrase, and they spread only leftward
  - **Scaling** of highs/lows:
    - Hs (and Ls) within an intermediate phrase are scaled to more-or-less the same heights
    - Downstep lowers the pitch ceiling until pitch reset (e.g., at a new ip), and requires a preceding H target within the same ip
    - There is no corresponding upstep

Some Questions

- **Empirical Question**: What types of intonational contours are found in rising AmE PQs?
- **Theoretical Question**: How do we model the attested variation in PQ intonation?

A Hunch

- **Hypothesis**: Perhaps some SPMs are meaningful / predictable
  - To probe this, we would need to understand the semantic/pragmatic context

Methods

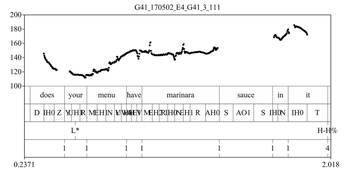
- Data collection:
  - 20 minutes of playing a modified version of the boardgame *Guess Who?*, in which players ask each other PQs
  - Speakers alternated between being told to play 'enthusiastically' and 'neutrally'
- Data preparation:
  - All questions were transcribed ([3]) and force-aligned ([5])
  - The data was annotated for a variety of semantico-pragmatic factors:
    - Aboutness: if the question is about menu item, number, game rules, etc.
    - Semantic force: if the question is information-seeking, confirmatory, etc.
    - Word order: if the question is polar, wh, alternative, etc.
  - ...and was annotated as best as possible with MAE_ToBI ([2])

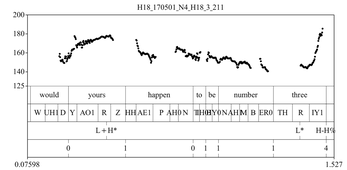
Data Set

- We collected 2,100 questions from 20 speakers, and annotated 1,592 of them
  - 19 speakers completely annotated
  - One speaker comprised half of the data, and was only partially annotated
- Only final-rise questions without disfluencies ($n = 857$) were analyzed in this study
  - (There were 1,011 final rises, of which 154 were disfluent)

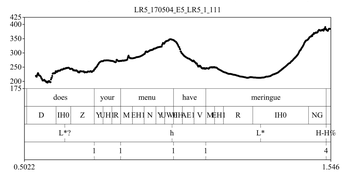
Some Contours for Polar Questions

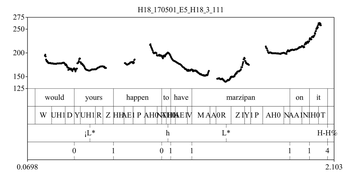
- 'Canonical' L* H-H% PQs (cf. [6]):



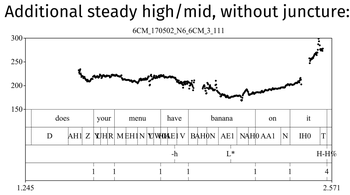


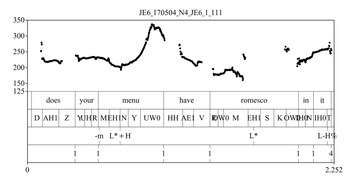
- Additional pointwise high/mid, without prominence:



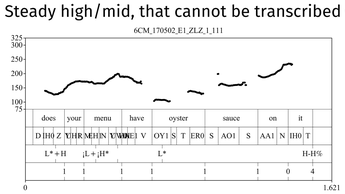


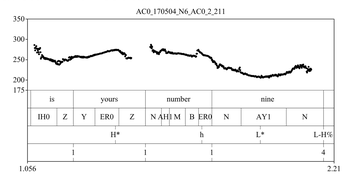
  - **Alignment**: a L* pitch accent can have a pitch peaks quite a bit later
- Additional steady high/mid, without juncture:





  - **Directionality**: these cover tones appear to be spreading from the left
- Steady high/mid, that cannot be transcribed with standard ToBI labels:





  - **Scaling**: Is there a ¡L* and/or !H* without a preceding high (or a dynamic pitch range uncoupled from 3-level breaks)?

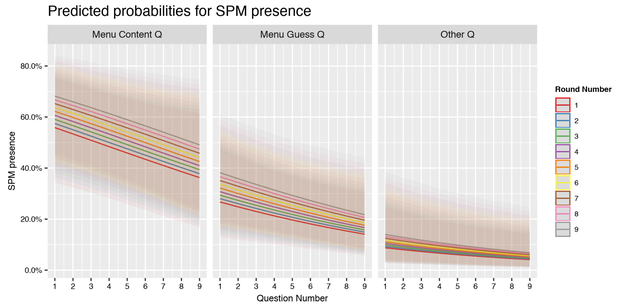
Data Analysis

- We did model comparison of linear mixed effects models in R, using the `lme4` package
  - Random effects: speaker
  - We sequentially introduced fixed effects that were of interest to us and then compared them using the `anova()` function
  - Using the `AIC()` function resulted in the same selection for best model

Results

- SPMs were only somewhat likely, once speaker variation was accounted for
- **Emotional state**: SPMs were more likely in the 'enthusiastic' condition
- **Question number**: SPMs became less likely over the course of a round
- **Round number**: SPMs became more likely in later rounds
- **Aboutness**: SPMs were more likely in menu-content questions than menu-guess questions
  - (Other questions were much less likely to have an SPM)
  - Intriguingly, there was a significant interaction of emotional state and round number: SPMs were less likely to appear in later parts of the task for rounds where speakers were instructed to be excited
    - This might be simply accounting for the effects of fatigue

Visualizing Some Results



Possible Sources for SPMs

- Paralinguistic effects support analyzing **SPMs as sub-phonemic**
  - There was a significant effect of emotional state
    - Perhaps non-time-locked SPMs are essentially cue strengtheners, employed when speakers are more emotional
  - There was a significant effect of question number
    - Within a round, perhaps there is less of a need to signal L*, or there was less emotional involvement
  - There was a significant effect of round number
    - Perhaps as rounds went on, participants spoke more fluenty (more SPMs)
- Semantic/Pragmatic effects support analyzing **SPMs as phonologically derived**
  - Increase in question number mirrors an enrichment of knowledge / discourse structure
  - Question aboutness is coextensive with particular semantic/pragmatic structures
    - Menu content questions are likely to be 'genuine' information seeking, while menu guesses take the form of a PQ but do not seek out information
      - Recall that all PQs analyzed here had final-rises!
    - Perhaps there are phonological categories / prosodic structures / intonational processes that we need to posit
      - /H+L*/? /HL*/? /-H/? Upstep? Range adjustments within intermediate phrases?

Conclusions

- We regularly observe **unpredicted pitch movements**
  - And different types! Within AmE PQs with final rises!
- There are **multiple contributing factors for SPMs**
  - Some contributing factors are **emotional/paralinguistic**
  - Others seem to be **phonological/semantic-pragmatic**
- **Further investigation required** to understand SPMs fully
  - Are certain SPMs conditioned differently from others?
  - How do people understand SPMs, in perception?

[1] D. Bates et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1. [2] M. E. Beckman, J. Hirschberg, and S. Shattuck-Hufnagel (2005). "The Original ToBI System and the Evolution of the ToBI Framework". In: *Prosodic Typology: The Phonology of Intonation and Phrasing*. Ed. by S.-A. Jun. Oxford: Oxford University Press. [3] P. Boersma and D. Weenink (2014). *Praat: doing phonetics by computer [Computer program]*. Available at http://www.praat.org/. [4] D. Lüdecke (2018). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.6.0. [5] M. McAuliffe et al. (2017). *Montreal Forced Aligner (version 0.9.0)*. Computer program. [6] J. Pierrehumbert and J. Hirschberg (1990). "The Meaning of Intonational Contours in the Interpretation of Discourse". In: *Intentions in Communication*. Ed. by P. R. Cohen, J. L. Morgan, and M. E. Pollack. Cambridge, MA: MIT Press. [7] R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.